

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-307849

(43) 公開日 平成10年(1998)11月17日

(51) Int.Cl.⁶
G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/403 3 2 0 D
15/40 3 7 0 A
15/401 3 1 0 C
15/403 3 2 0 A

審査請求 未請求 請求項の数9 O L (全 9 頁)

(21) 出願番号 特願平10-51694

(22) 出願日 平成10年(1998) 3 月 4 日

(31) 優先権主張番号 特願平9-52170

(32) 優先日 平 9 (1997) 3 月 6 日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000102728

株式会社エヌ・ティ・ティ・データ
東京都江東区豊洲三丁目 3 番 3 号

(72) 発明者 中島 浩之

東京都江東区豊洲三丁目 3 番 3 号 エヌ・
ティ・ティ・データ通信株式会社内

(72) 発明者 木谷 強

東京都江東区豊洲三丁目 3 番 3 号 エヌ・
ティ・ティ・データ通信株式会社内

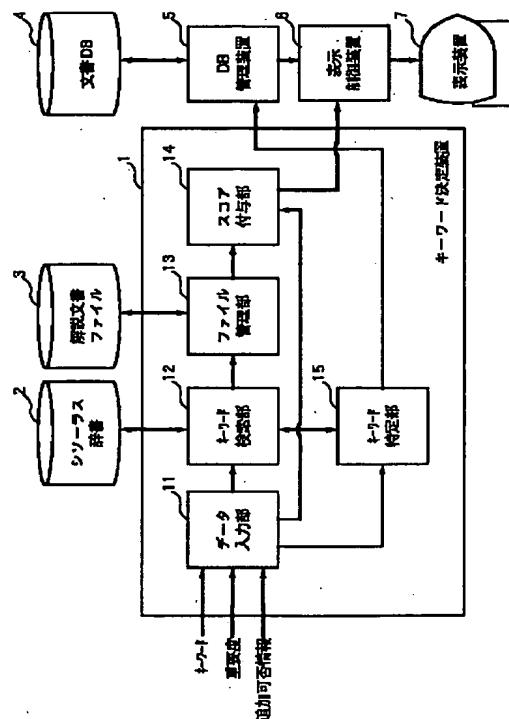
(74) 代理人 弁理士 鈴木 正剛

(54) 【発明の名称】 検索用のキーワードの決定方法及び装置、文書検索装置、記録媒体

(57) 【要約】

【課題】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置において、真に必要なキーワードのみを特定できるようにする。

【解決手段】 キーワード決定装置 1 において、入力キーワードに追加すべきキーワード候補をシソーラス辞書 2 から索出するとともに、そのキーワード候補の解説文書を解説文書ファイル 3 から索出する。そして、解説文書等に入力キーワードやキーワード候補との関連度合いに応じたスコアを付与し、これらを表示装置 7 へ表示させて、操作者に当該キーワード候補の追加可否判断を促す。操作者からの判断結果を解読し、そのキーワード候補が追加不要であれば、入力キーワードのみを検索用のキーワードとする。一方、キーワード候補の追加が必要である場合は、そのキーワード候補を検索用のキーワードに追加する。



(2)

【特許請求の範囲】

【請求項1】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置に予めキーワード候補と当該キーワード候補の電子解説情報とをリンクさせて保持しておき、

任意の入力キーワードを受け付けたときに、前記文書検索装置上で、当該入力キーワードに追加すべき他のキーワード候補とそのキーワード候補にリンクする電子解説情報とを取得する処理と、取得したキーワード候補を電子解説情報と共に操作者が認識可能な形態で提示する処理と、該提示に応じて操作者より入力された当該キーワード候補の追加可否情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定する処理とを少なくともこの順に実行することを特徴とするキーワード決定方法。

【請求項2】 前記電子解説情報に対して前記入力キーワード及びキーワード候補との関連度合いを表すスコアを付与し、前記キーワード候補及び電子解説情報を提示する際に、前記スコアに応じて異なる提示形態を形成することを特徴とする請求項1記載のキーワード決定方法。

【請求項3】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置において一または複数のキーワードを決定する装置であって、入力キーワードに対して追加すべきキーワード候補とそのキーワード候補に対応する電子解説情報とを取得するキーワード情報取得手段と、取得した前記キーワード候補及び電子解説情報を前記文書画像検索装置が認識可能な情報形態に加工するとともに、前記文書検索装置に操作者に対して当該キーワード候補の追加可否情報の入力を促す提示形態を形成させる提示制御手段と、を有することを特徴とするキーワード決定装置。

【請求項4】 前記取得した電子解説情報に対して前記入力キーワード及びキーワード候補との関連度合いを表すスコアを付与するスコア付与手段をさらに有し、前記提示制御手段は、前記キーワード候補、及び／または、そのキーワード候補の電子解説情報を、前記スコア付与手段が付与したスコアに応じて異なる形態で文書検索装置に提示させるように構成されていることを特徴とする請求項3記載のキーワード決定装置。

【請求項5】 前記文書検索装置による提示に応じて操作者より入力されたキーワード候補の追加可否情報を解読し、解読結果に応じて当該キーワード候補をシソーラス検索用のキーワードとして追加するか否かを決定する手段をさらに有することを特徴とする請求項3または4記載のキーワード決定装置。

【請求項6】 決定されたシソーラス検索用のキーワードの各々に検索目的に応じた重み情報を付加する手段をさらに有することを特徴とする請求項5記載のキーワ

ード決定装置。

【請求項7】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置であって、個々のキーワードに対して追加すべきキーワード候補を記録したシソーラス辞書と、キーワード候補毎にそのキーワード候補の電子解説情報を記録した情報ファイルと、前記シソーラス辞書及び情報ファイルの記録情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定するキーワード決定装置とを含んで成り、

- 10 該キーワード決定装置が、入力キーワードに追加すべきキーワード候補を前記シソーラス辞書より取得するとともに取得したキーワード候補の電子解説情報を前記情報ファイルより取得するキーワード情報取得手段と、取得したキーワード候補及び電子解説情報を操作者が認識可能な形態で提示する情報提示手段と、該情報提示手段による提示に応じて操作者より入力されたキーワード候補の追加可否情報を解読し、解読結果に応じて当該キーワード候補をシソーラス検索用のキーワードとして追加するか否かを決定する手段とを有することを特徴とする文書検索装置。

【請求項8】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置をコンピュータ上で実現するためのプログラムがコンピュータ読み取り可能な形態で記録された記録媒体であって、前記プログラムが、少なくとも、予めキーワード候補と当該キーワード候補の電子解説情報とをリンクさせて保持する処理、任意の入力キーワードを受け付けたときに、当該入力キーワードに追加すべき他のキーワード候補とそのキーワード候補にリンクする電子解説情報とを取得する処理、取得したキーワード候補を電子解説情報と共に操作者が認識可能な形態で提示する処理、該提示に応じて操作者より入力された当該キーワード候補の追加可否情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定する処理を、前記コンピュータに実行させるものであることを特徴とする記録媒体。

- 40 【請求項9】 相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置をコンピュータ上で実現するためのプログラムがコンピュータ読み取り可能な形態で記録された記録媒体であって、前記プログラムが、少なくとも、個々のキーワードに対して追加すべきキーワード候補を記録したシソーラス辞書及びキーワード候補毎にそのキーワード候補の電子解説情報を記録した情報ファイルを前記コンピュータに形成させ、さらに、入力キーワードに追加すべきキーワード候補を前記シソーラス辞書より取得するとともに、このキーワード候補
- 50 に対応する電子解説情報を前記情報ファイルより取得す

3

る処理、
 取得したキーワード候補及び電子解説情報を操作者が認識可能な形態で提示する処理、
 該提示に応じて操作者より入力されたキーワード候補の追加可否情報を解読し、解読結果に応じて当該キーワード候補をシソーラス検索用のキーワードとして追加するか否かを決定する処理、
 を前記コンピュータに実行させるものであることを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置に関し、特に、検索時に入力されたキーワード（入力キーワード）に追加すべき他のキーワード（キーワード候補）が、検索目的に適合したキーワードかどうかを決定して不要な検索処理を回避するための技術に関する。

【0002】

【従来の技術】予めデータベース化された複数の電子文書から入力キーワードとの関連性がより高い電子文書を検索する装置として、入力キーワード毎に、その入力キーワードに関連する他のキーワードを追加し、相互に関連するこれらの複数のキーワードを用いてシソーラス検索を行う文書検索装置が知られている。この種の文書検索装置はコンピュータによって実現されるもので、図4の機能ブロック図に示すように、キーワード決定装置41、シソーラス辞書42、文書データベース（文書DB）43、データベース（DB）管理装置44、表示制御装置45、及び表示装置46を含んで構成される。

【0003】シソーラス辞書42は、キーワード毎に、追加すべき関連キーワード（追加キーワード）を記録したものであり、文書DB43は、検索対象となる複数の電子文書を格納したものである。キーワード決定装置41は、キーワードや検索目的に応じた重み情報、すなわち重要度を入力するデータ入力部47、入力キーワードに関連する追加キーワードをシソーラス辞書42より索出するキーワード検索部48、及び入力キーワード及び追加キーワードにそれぞれ重要度を設定して検索に用いる複数のキーワードを特定するキーワード特定部49の機能を備えるものである。特定されたキーワードは、DB管理装置44に送られる。DB管理装置44は、複数のキーワードと各キーワードに設定された重要度に基づいて文書DB43より該当する電子文書を索出し、表示制御装置45を通じて電子文書を表示装置46に表示させる。

【0004】上記構成の文書検索装置において、入力キーワード及び重要度が取得されてから、シソーラス検索に用いる複数のキーワードが特定されるまでの過程を図5に示す。図5中、符号42、47、49は、図4にお

(3)

4

けるキーワード決定装置41の該当機能ブロックを表している。

【0005】図5を参照すると、入力キーワード「kwd1」、「kwd2」、「kwd3」に関連する追加キーワードとして、それぞれシソーラス辞書42から「kwd8」、「kwd6」、「kwd7」が索出されている。また、入力キーワード「kwd1」、「kwd2」、「kwd3」の重要度として、それぞれ「2」、「10」、「6」が設定されている。追加キーワード

10 「kwd8」、「kwd6」、「kwd7」の重要度については、これらが入力キーワードから二次的に導かれていることを考慮して、入力キーワード「kwd1」、「kwd2」、「kwd3」に設定された重要度の値の1/2の値、つまり、「1」、「5」、「3」が設定されている。DB管理装置44は、上記複数のキーワードと重要度とを用いたシソーラス検索（文書検索）を行い、該当する電子文書を文書DB43より索出して、これを表示制御装置45へ送る。表示制御装置45は、この電子文書を表示装置46に画面表示する。

20 【0006】このように、従来の文書検索装置では、4つの入力キーワード「kwd1」～「kwd4」だけでなく、追加キーワード「kwd6」～「kwd9」をも加えた8つのキーワードによってシソーラス検索を行うことで、広範囲な検索を可能にするとともに、検索時に各キーワードの重要度をも考慮することで、検索目的に合致した電子文書の取得を可能にしている。

【0007】

【発明が解決しようとする課題】上述のように、従来の文書検索装置では、追加キーワードを全てシソーラス検索用のキーワードとして用いている。しかし、追加キーワードの中には、操作者にとって必ずしも重要でなかったり、全く不要なものが含まれる場合があり、このような不要な追加キーワードに基づくシソーラス検索は無駄な処理となる。そのため、従来の文書検索装置では、検索時間の短縮化が図れないばかりでなく、操作者が真に望む電子文書とは全く異なる電子文書が索出されてしまう場合が多かった。例えば、パーソナルコンピュータの製品名である「マッキントッシュ」の略語である「マック」を入力キーワードとした場合に、追加キーワードとして、ファーストフードのチェーン店である「マクドナルド」が抽出されることがある。この場合、コンピュータとは全く関係のない、食品業界関係の情報を記述した電子文書が検索されてしまう。

【0008】そこで本発明の課題は、相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置において、入力キーワードに他の関連キーワードを追加する際に、その関連キーワードが操作者にとって真に検索に必要なキーワードかどうかを判定して無駄な検索処理を回避することにある。本発明の他の課題は、上記文書検索処理をコンピュータで実現するための記録媒

50

5

体を提供することにある。

【0009】

【課題を解決するための手段】上記課題を解決するため、本発明は、相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置において適用される、改良されたキーワード決定方法を提供する。この方法は、文書検索装置に予めキーワード候補と当該キーワード候補の電子解説情報とをリンクさせて保持しておき、任意の入力キーワードを受け付けたときに、文書検索装置上で、当該入力キーワードに追加すべき他のキーワード候補とそのキーワード候補にリンクする電子解説情報とを取得する処理と、取得したキーワード候補を電子解説情報と共に操作者が認識可能な形態で提示する処理と、該提示に応じて操作者より入力された当該キーワード候補の追加可否情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定する処理とを少なくともこの順に実行することを特徴とする。電子解説情報は、例えばそのキーワード候補を含む周辺知識を記録した解説文書であり、それが提示されたときに操作者がキーワード候補の意味内容を把握できる内容のものである。

【0010】このように、文書検索装置側から操作者にキーワード候補と電子解説情報を併せて提示することで、操作者に対してキーワード候補を追加すべきかどうかの判断の便宜を図り、操作者による判断結果に応じてキーワードの追加可否を決定するようにして、不要なキーワードに基づく検索処理を回避する。

【0011】本発明は、また、上記方法の実施に適した文書検索装置をも提供する。本発明の文書検索装置は、相互に関連する複数のキーワードによるシソーラス検索が可能な装置であって、個々のキーワードに対して追加すべきキーワード候補を記録したシソーラス辞書と、キーワード候補毎にそのキーワード候補の電子解説情報を記録した情報ファイルと、前記シソーラス辞書及び情報ファイルの記録情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定するキーワード決定装置とを含んで構成される。

【0012】キーワード決定装置は、以下のような構成要素の全てあるいは一部の組み合わせから成り、文書検索装置に、操作者に対して前記キーワード候補の追加可否情報の入力を促す提示形態を形成させるとともに、操作者からの追加可否情報に基づいて合理的な検索を行うためのキーワードを決定する。

(1) 入力キーワードに対して追加すべきキーワード候補とそのキーワード候補の電子解説情報とを取得するキーワード情報取得手段。

(2) キーワード情報取得手段で取得したキーワード候補または電子解説情報に対して前記入力キーワードとの関連度合いを表すスコアを付与するスコア付与手段。

(3) 前記キーワード情報取得手段で取得したキーワー

(4)

6

ド候補をその電子解説情報と共に操作者が認識可能な形態で提示させる提示制御手段。なお、前述のようにスコアが電子解説情報に付与される場合、この提示制御手段は、スコアに応じて異なる形態でキーワード候補や電子解説情報を提示させるように構成する。

(4) 前記提示制御手段による提示に応じて操作者より入力されたキーワード候補の追加可否情報を解読し、解読結果に応じて当該キーワード候補をシソーラス検索用のキーワードとして追加するか否かを決定する手段。

10 (5) 決定されたシソーラス検索用のキーワードの各々に検索目的に応じた重み情報を付加する手段。

【0013】上記他の課題を解決する本発明の記録媒体は、相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置をコンピュータ上で実現するためのプログラムがコンピュータ読み取り可能な形態で記録された記録媒体であって、前記プログラムが、少なくとも、予めキーワード候補と当該キーワード候補の電子解説情報とをリンクさせて保持する処理、任意の入力キーワードを受け付けたときに、当該入力キーワードに追加すべき他のキーワード候補とそのキーワード候補にリンクする電子解説情報とを取得する処理、取得したキーワード候補を電子解説情報と共に操作者が認識可能な形態で提示する処理、該提示に応じて操作者より入力された当該キーワード候補の追加可否情報に基づいてシソーラス検索に用いる一または複数のキーワードを決定する処理を、前記コンピュータに実行させるものである。

【0014】本発明の他の記録媒体は、相互に関連する複数のキーワードによるシソーラス検索が可能な文書検索装置をコンピュータ上で実現するためのプログラムがコンピュータ読み取り可能な形態で記録された記録媒体であって、前記プログラムが、少なくとも、個々のキーワードに対して追加すべきキーワード候補を記録したシソーラス辞書及びキーワード候補毎にそのキーワード候補の電子解説情報を記録した情報ファイルを前記コンピュータに形成させ、さらに、入力キーワードに追加すべきキーワード候補を前記シソーラス辞書より取得するとともに、このキーワード候補に対応する電子解説情報を前記情報ファイルより取得する処理、取得したキーワード候補及び電子解説情報を操作者が認識可能な形態で提示する処理、該提示に応じて操作者より入力されたキーワード候補の追加可否情報を解読し、解読結果に応じて当該キーワード候補をシソーラス検索用のキーワードとして追加するか否かを決定する処理、を前記コンピュータに実行させるものである。

【0015】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を詳細に説明する。図1は、本発明の一実施形態に係る文書検索装置の機能ブロック図である。この文書検索装置は、パーソナルコンピュータやワークステーション等のコンピュータで実現される。すなわち、通信

50

(5)

7 ネットワークを通じて入力され、あるいはCD-ROM (コンパクトディスク型ROM) やFD (フレキシブルディスク) 等の可搬性メディアから読み出したプログラムを図示しないコンピュータメモリにインストールし、該コンピュータが上記プログラムを随時読み出して実行することによって形成される、キーワード決定装置1、シソーラス辞書2、情報ファイルである解説文書ファイル3、文書DB4、DB管理装置5、表示制御装置6と、CRTあるいは液晶パネルからなる表示装置7とを含んで構成される。

【0016】シソーラス辞書3、文書DB4、DB管理装置5、表示装置7は、それぞれ図4に示した従来装置のシソーラス辞書42、文書DB43、DB管理装置44、表示装置46と同一機能のものである。解説文書ファイル3には、キーワード候補毎にそのキーワード候補の意味内容、周辺知識、その他関連する情報等を記述した解説文書(電子解説情報)が記録されている。この場合、解説文書は、一つのキーワード候補について複数のものを用意しておき、各々の関連度合いに応じて特定のものを選択できるようにしておく。

【0017】キーワード決定装置1は、シソーラス検索に用いる合理的なキーワードを決定するために、データ入力部11、キーワード検索部12、ファイル管理部13、スコア付与部14、及びキーワード特定部15の機能を備えて構成される。データ入力部11は、入力キーワードやその重要度、キーワード候補の追加可否情報、その他のデータの入力を受け付けるものであり、キーワード検索部12は、入力キーワードに関連する他のキーワードをシソーラス辞書2から索出してこれを追加すべきキーワード候補として一時的に保持するものである。

【0018】ファイル管理部13はキーワード候補とそのキーワード候補の解説文書の記録領域とをリンクさせて管理するものである。具体的には、あるキーワード候補の解説文書が解説文書ファイル3のどの記録領域(コード番号)に存在するかを表すテーブルを有しており、キーワード候補が特定されたとき、このキーワード候補についての解説文書をコード番号を通じて解説文書ファイル3より取得できるようになっている。

【0019】スコア付与部14は、個々の解説文書に対して入力キーワードやキーワード候補との関連度合いを表すスコア(評価値)を付与するものである。このスコアは、同一のキーワード候補について複数の解説文書があるときの選択順序の特定と、表示制御装置6において操作者によるキーワード候補の追加可否判断を容易にするための表示形態の形成時に用いられるものである。スコアの付与については種々の手法を採用することができる。例えば、公知のTF/IDF法(「Introduction to Modern Information Retrieval」, Gerard Salton他著, McGraw-Hill Publishing Company刊)や、本出願人により先に提案された特願平8-110870号明細書「類似文

8

書検索装置」に開示されているスコア付与手法等を用いることができる。

【0020】キーワード特定部15は、操作者より入力された当該キーワード候補の追加可否情報を解読し、解読の結果、追加不要(操作者がそのキーワード候補を採用しない)と判定した場合は、そのキーワード候補を破棄(削除)して入力キーワードのみを検索用のキーワードとする。一方、解読の結果、追加が必要(操作者がそのキーワード候補を採用する)と判定した場合は、その
10 キーワード候補を入力キーワードに追加する。また、入力キーワードと追加されたキーワードの重要度をも決定し、これらを各キーワード毎に設定してDB管理装置5へ出力する。入力キーワードの重要度は、キーワードと共に入力されたものをそのまま用いる。追加されたキーワードの重要度については、入力キーワードの重要度の値をそのまま用いてもよいし、入力キーワードの重要度の値に係数(例えば、従来装置の場合と同様の0.5)を乗算した値を用いてもよい。この例では、後者の値を用いる。

20 【0021】表示制御装置6は、スコアが付与された解説文書とキーワード候補とをキーワード決定装置1より受領し、キーワード候補とその解説文書とを表示装置7に同時に表示させるための制御を行うものである。表示させる形態は、例えば、キーワード候補と解説文書の一部(例えばタイトルのみ、タイトルと著者、あるいは要約のみ)を表示させる形態、キーワード候補と解説文書の全文を表示させる形態、キーワード候補のみを点滅あるいは他の部分と異なる色を用いて強調表示させる形態等、操作者が注意を喚起し得る種々の形態を任意に選択
30 できるようにする。また、操作者がキーワード候補の追加可否の判断を容易に行えるようにするため、解説文書に付与された上記スコアを識別し、スコアの高い解説文書ほど画面上部に位置させたり、あるいは大きな文字、インパクトの強い色等で強調表示させるようにしてもよい。また、解説文書をスタック形式で表示させ、スコアの高い解説文書ほどスタックの上位に表示させるようにしてもよい。

【0022】次に、図2及び図3を参照して、本実施形態による文書検索装置の動作を具体的に説明する。図2
40 は、本実施形態の文書検索装置において用いる入力キーワード、キーワード候補、解説文書、スコア、重要度の具体例を示す説明図であり、図3は、主としてキーワード決定装置1におけるキーワード特定処理の手順を示すフローチャートである。図2において、符号2, 3, 7, 11, 13, 15は、それぞれ図1におけるキーワード決定装置1の該当機能ブロックを表している。

【0023】いま、図2に示すように「エイズ」、「薬害」、「感染症」の3つの入力キーワードがデータ入力部11に入力されたとする(ステップS101)。キーワード検索部12は、まず、シソーラス辞書2に、「エ

50

9

イズ」について追加すべきキーワード候補が記録されているかどうかを検索する（ステップS102）。この例では、「エイズ」に対するキーワード候補として「血液製剤」がシソーラス辞書2に記録されているので、キーワード検索部12は、この「血液製剤」を索出してファイル管理部13へ送る。

【0024】ファイル管理部13は、キーワード候補である「血液製剤」からそれが記録されている領域のコード番号「4000」を特定する。そして、このコード番号「4000」に対応した解説文書ファイル3の記録領域から解説文書「1」～「3」を索出し、これをスコア付与部14へ送る（ステップS103）。

【0025】スコア付与部14は、入力キーワードとその重要度、及びキーワード候補に基づいて、解説文書「1」～「3」に対してそれぞれスコアを付与する（ステップS104）。解説文書「1」には、重要度「2」の「エイズ」と、重要度「10」の「薬害」と、重要度「6」の「感染症」とが含まれているため、スコアは他の解説文書「2」、「3」よりも高くなっている。スコアが付与された解説文書「1」～「3」は、キーワード候補である「血液製剤」と共に表示制御装置6へ送られる。表示制御装置6は、「血液製剤」と各解説文書「1」～「3」とを共に表示装置7に表示させる（ステップS105）。

【0026】「血液製剤」等の表示に基づいて操作者からキーワード候補の追加可否情報が入力されると、キーワード特定部15は、この追加可否情報を解読する。そして解読の結果、「血液製剤」の追加が選択されなかった場合（ステップS106：No）、キーワード特定部15は、自己あるいはキーワード検索部12に保持されている「血液製剤」を削除する。一方、「血液製剤」の追加が選択された場合（ステップS106：Yes）は、

「血液製剤」の重要度を設定する（ステップS108）。この実施形態では、もとの入力キーワードの重要度の0.5倍の値をキーワード候補の重要度としており、入力キーワード「エイズ」の重要度が「2」であることから、「血液製剤」の重要度を「1」に設定する。

【0027】次のキーワード候補がある場合はステップS102の処理に戻り、次のキーワード候補がない場合は検索用のキーワードを決定してDB管理装置5へ出力する（ステップS109）。つまり、操作者が「血液製剤」の追加を選択しなかった場合は「エイズ」、「薬害」、「感染症」の3つの入力キーワードのみを検索用のキーワードとして決定し、一方「血液製剤」の追加を選択した場合はこれを検索用のキーワードに追加して、DB管理装置5へ出力する。DB管理装置5は、決定された検索用のキーワードをこれらの重要度をも考慮したうえで、電子文書の検索を行う。

【0028】このように、本実施形態の文書検索装置では、キーワード追加装置1が、入力キーワードに基づい

(6)

10

て索出されたキーワード候補の追加可否を操作者に判断させた後に検索用のキーワードを決定するので、不要なキーワードによる検索処理を回避することができる。また、キーワード候補と共にその解説文書をも表示させるようにしたので、操作者がキーワード候補の意味内容やその関連知識を別途調べなくとも、容易かつ的確にそのキーワード候補が検索に必要なかどうかを判断できるようになり、検索に要する時間の短縮化を図ることもできる。さらに、各キーワードに付与された重要度をも考慮したうえでシソーラス検索を行うことができるので、検索精度を高めることもできる。

【0029】なお、本実施形態では、キーワード候補を一つずつ処理する場合の例について説明したが、複数のキーワード候補について同時に上記処理を行うようにしてもよい。また、この実施形態では、一つの入力キーワードに対するキーワード候補の数は一つとして説明したが、入力キーワードに対するキーワード候補の数は任意に決定することができる。

【0030】さらに、本実施形態では、操作者に対するキーワード候補や電子解説情報の提示の形態として、表示装置6を用いた画面表示の例について説明したが、画面表示に代えて、あるいは画面表示と共に、音声合成装置や既存のテキストリーダ等を用いて音声の形で提示するようにしてもよい。キーワード候補等を音声で出力する場合、上述のスコアは、例えば音量の強弱や音色を変化させるためのパラメータとして用いることができる。

【0031】

【発明の効果】以上の説明から明らかなように、本発明によれば、キーワード候補がその電子解説情報と共に操作者に提示されるので、操作者は、そのキーワード候補の追加が必要であるか否かの判断を容易かつ正確に行うことができるようになる。

【0032】また、操作者によって必要であると判断したキーワード候補のみが追加され、余計なキーワード候補は検索に用いられなくなるので、無駄な検索処理が回避され、従来よりも一層高速かつ正確な文書検索が可能となる。

【0033】更に、電子解説情報に、入力キーワードとの関連度合いに応じたスコアが付与されて提示されるので、操作者によるキーワード候補の追加可否判断がより一層容易かつ正確になる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る文書検索装置の機能ブロック図。

【図2】本実施形態の文書検索装置における入力キーワード、キーワード候補、解説文書、スコア、重要度の例と、これらに基づいて検索用のキーワードが特定されるまでの過程を示す説明図。

【図3】本実施形態の文書検索装置によるキーワード特定処理の手順を示すフローチャート。

(7)

11

12

【図4】従来の文書検索装置の機能ブロック図。

【図5】従来の文書検索装置における入力キーワード、追加キーワード、重要度の例と、検索用のキーワードが特定されるまでの過程を示す説明図。

【符号の説明】

- 1 キーワード決定装置
- 2 シソーラス辞書
- 3 解説文書ファイル
- 4 文書DB (データベース)

5 DB管理装置

6 表示制御装置

7 表示装置

11 データ入力部

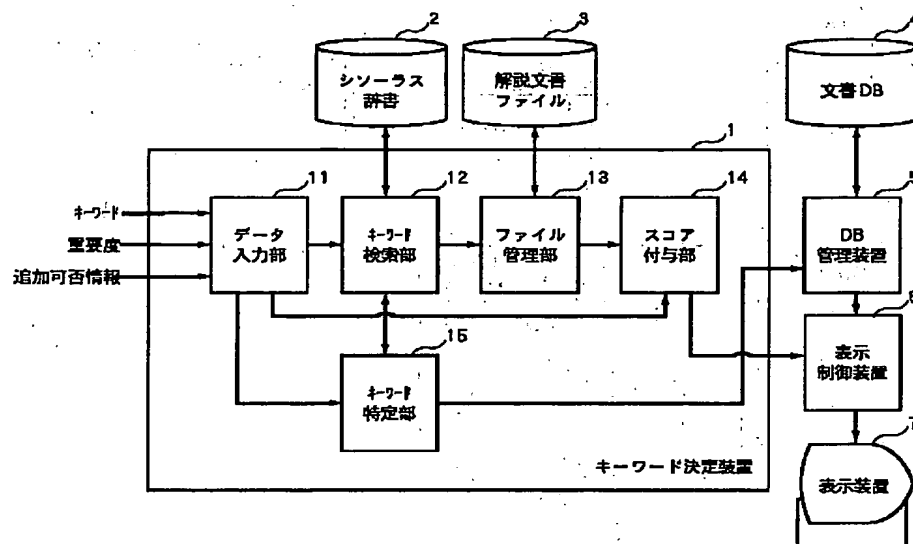
12 キーワード検索部

13 ファイル管理部

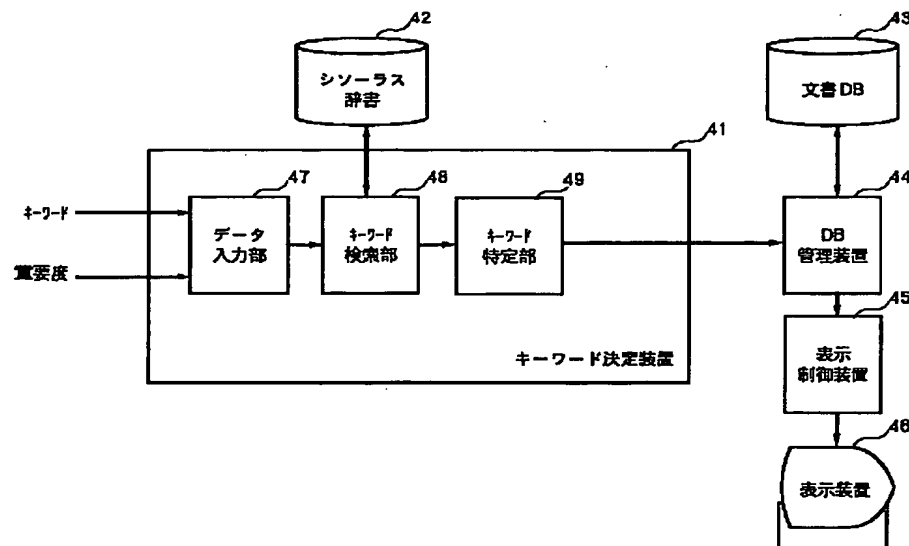
14 スコア付与部

15 キーワード特定部

【図1】

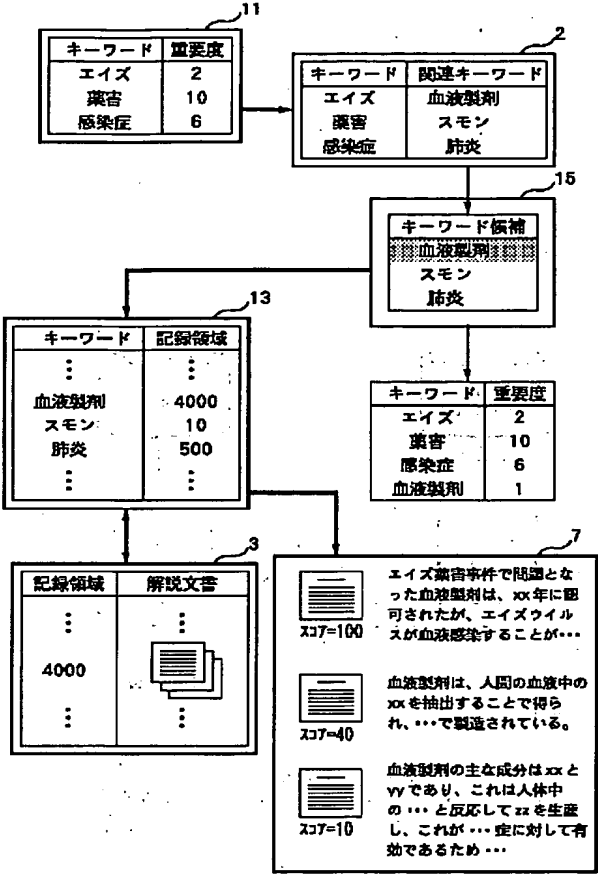


【図4】

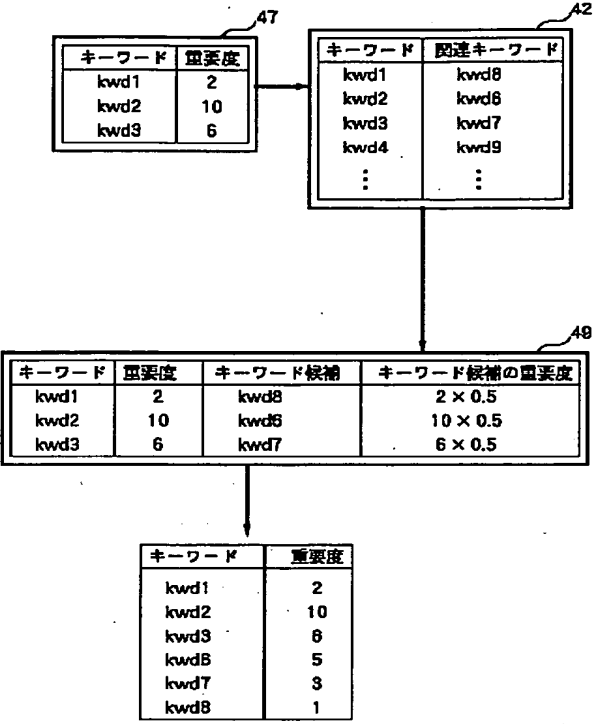


(8)

【図2】

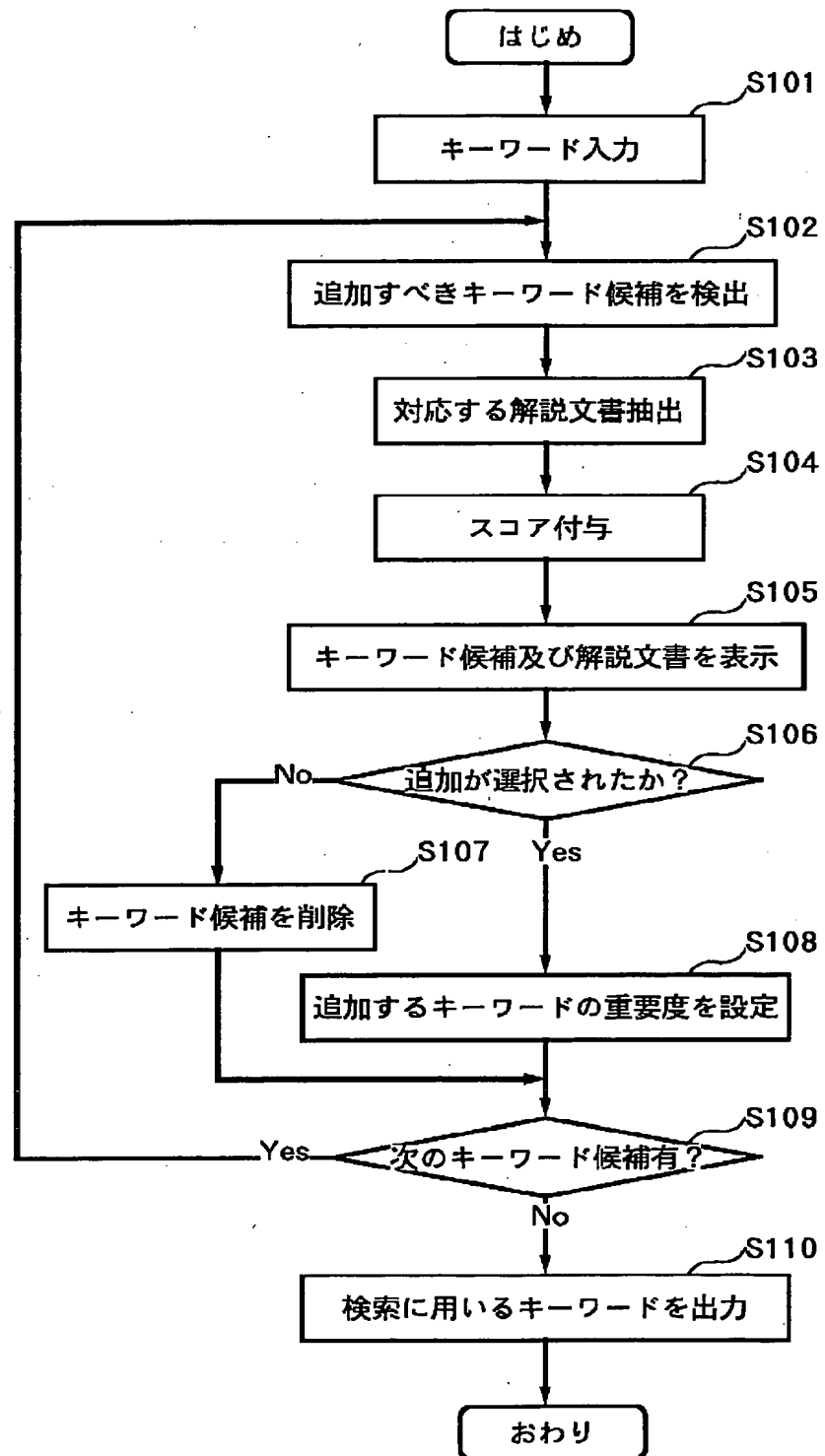


【図5】



(9)

【図3】



PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-307849

(43)Date of publication of application : 17.11.1998

(51)Int.Cl.

G06F 17/30

(21)Application number : 10-051694

(71)Applicant : N T T DATA:KK

(22)Date of filing : 04.03.1998

(72)Inventor : NAKAJIMA HIROYUKI
KITANI TSUYOSHI

(30)Priority

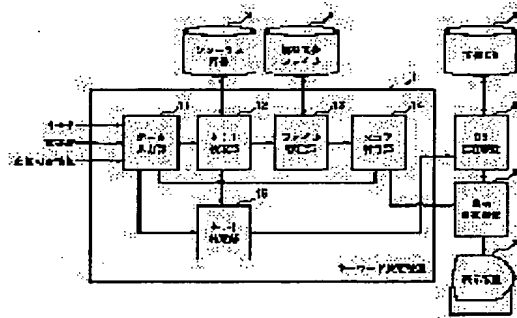
Priority number : 09 52170 Priority date : 06.03.1997 Priority country : JP

(54) RETRIEVING KEYWORD DETERMINING METHOD, ITS DEVICE, DOCUMENT RETRIEVING DEVICE, AND RECORDING MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To specify only a really necessary keyword in a document retrieving device capable of retrieving a thesaurus by plural mutually relational keywords.

SOLUTION: A keyword determination device 1 retrieves a keyword candidate to be added to an input keyword from a thesaurus dictionary 2 and retrieves an instruction document for the keyword candidate from an instruction document file 3. A score corresponding to the degree of relation with the input keyword and the keyword candidate is applied to the document or the like and displayed on a display device 7 to suggest judgement whether a keyword candidate is to be added or not to an operator. When the addition of the keyword candidate is unnecessary after decoding the judged result from the operator, only the input keyword is selected as a keyword for retrieval. When the addition of the keyword candidate is necessary, the candidate is added to the retrieving keyword.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] When the electronic description information of a keyword candidate and the keyword candidate concerned is made to link to the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible beforehand, it holds and arbitrary input keywords are received The processing which acquires the electronic description information linked to other keyword candidate who should add to the input keyword concerned, and its keyword candidate on the aforementioned document-retrieval equipment, The processing which presents the keyword candidate who acquired with electronic description information with the form which an operator can recognize, The keyword determination method characterized by performing at least processing which determines 1 or two or more keywords which are used for thesaurus reference based on the additional propriety information of the keyword candidate concerned that it was inputted by the operator according to this presentation in this order.

[Claim 2] The keyword determination method according to claim 1 characterized by forming a different presentation form according to the aforementioned score in case the score which expresses a related degree with the aforementioned input keyword and a keyword candidate to the aforementioned electronic description information is given and the aforementioned keyword candidate and electronic description information are shown.

[Claim 3] Equipment which determines 1 or two or more keywords in the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually characterized by providing the following is possible A keyword information acquisition means to acquire the electronic description information corresponding to the keyword candidate who should add to an input keyword, and its keyword candidate Presentation control means in which the presentation form which urges the input of the additional propriety information of the keyword candidate concerned to the aforementioned document-retrieval equipment to an operator is made to form while processing the aforementioned keyword candidate and electronic description information which were acquired into the information form which can recognize the aforementioned document picture reference equipment

[Claim 4] It is keyword determination equipment according to claim 3 carry out that have further a score grant means give the score which expresses a related degree with the aforementioned input keyword and a keyword candidate to the electronic description information which carried out [aforementioned] acquisition, and the aforementioned presentation control means are constituted so that the aforementioned score grant means may make the aforementioned keyword candidate and/or the keyword candidate's electronic description information show document-retrieval equipment with a different form according to the given score as the feature.

[Claim 5] Keyword determination equipment according to claim 3 or 4 characterized by having further a means to determine whether to decode a keyword candidate's additional propriety information that it was inputted by the operator according to presentation by the aforementioned document-retrieval equipment, and add the keyword candidate concerned as a keyword for thesaurus reference according to a decode result.

[Claim 6] Keyword determination equipment according to claim 5 characterized by having further a means to add the weight information according to the reference purpose to each of the determined keyword for thesaurus reference.

[Claim 7] the thesaurus reference by two or more keywords related mutually characterized by providing the following is possible -- it document-retrieval-equips The thesaurus dictionary which recorded the keyword candidate who should add to each keyword The information file which recorded the keyword candidate's electronic description information for every keyword candidate A keyword information acquisition means acquire a keyword candidate's electronic description information acquired while it changed including the keyword determination equipment which determines 1 or two or more keywords

which are used for thesaurus reference based on the recording information of the aforementioned thesaurus dictionary and an information file and this keyword determination equipment acquired the keyword candidate who should add to an input keyword from the aforementioned thesaurus dictionary from the aforementioned information file A means to determine whether to decode a keyword candidate's additional propriety information that it was inputted by the operator according to presentation by information presentation means to show the keyword candidate and electronic description information which were acquired with the form which an operator can recognize, and this information presentation means, and add the keyword candidate concerned as a keyword for thesaurus reference according to a decode result

[Claim 8] It is the record medium with which the program for realizing on a computer the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible was recorded with the form in which computer reading is possible. When the processing and the arbitrary input keywords which the aforementioned program makes link the electronic description information of a keyword candidate and the keyword candidate concerned beforehand, and holds at least are received The processing which acquires the electronic description information linked to other keyword candidate who should add to the input keyword concerned, and its keyword candidate, The processing which presents the keyword candidate who acquired with electronic description information with the form which an operator can recognize, The record medium characterized by being the thing which makes the aforementioned computer perform processing which determines 1 or two or more keywords which are used for thesaurus reference based on the additional propriety information of the keyword candidate concerned that it was inputted by the operator according to this presentation.

[Claim 9] It is the record medium with which the program for realizing on a computer the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible was recorded with the form in which computer reading is possible. The information file which recorded the keyword candidate's electronic description information for every thesaurus dictionary in which the aforementioned program recorded at least the keyword candidate who should add to each keyword, and keyword candidate is made to form in the aforementioned computer. further While acquiring the keyword candidate who should add to an input keyword from the aforementioned thesaurus dictionary The processing which acquires the electronic description information corresponding to this keyword candidate from the aforementioned information file, The processing which presents the keyword candidate and electronic description information which were acquired with the form which an operator can recognize, The record medium characterized by being the thing which makes the aforementioned computer perform processing which determines whether to decode a keyword candidate's additional propriety information that it was inputted by the operator according to this presentation, and add the keyword candidate concerned as a keyword for thesaurus reference according to a decode result.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] this invention relates to the technology for other keywords (keyword candidate) which should be added to the keyword (input keyword) especially inputted at the time of reference determining whether to be the keyword which suited the reference purpose, and avoiding unnecessary reference processing about the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible.

[0002]

[Description of the Prior Art] As equipment with which an electronic filing document with higher relevance with an input keyword is searched from two or more electronic filing documents database-ized beforehand, other keywords relevant to the input keyword are added for every input keyword, and the document-retrieval equipment which performs thesaurus reference using two or more of these keywords relevant to mutual is known. A computer realizes, and this kind of document-retrieval equipment is constituted including keyword determination equipment 41, the thesaurus dictionary 42, the document database (document DB) 43, database (DB) management equipment 44, a display controller 45, and display 46, as shown in the functional block diagram of drawing 4.

[0003] The thesaurus dictionary 42 records the related keyword (additional keyword) which should be added for every keyword, and a document DB43 stores two or more electronic filing documents used as the candidate for reference. Keyword determination equipment 41 is equipped with the function of the keyword specification section 49 which specifies two or more keywords which set significance as the data input section 47 which inputs the weight information according to the keyword or the reference purpose, i.e., significance, the retrieval-by-keyword section 48 which **** the additional keyword relevant to an input keyword from the thesaurus dictionary 42, an input keyword, and an additional keyword, respectively, and are used for reference. The specified keyword is sent to DB management equipment 44. DB management equipment 44 **** the electronic filing document which corresponds from a document DB43 based on the significance set as two or more keyword and each keyword, and displays an electronic filing document on display 46 through a display controller 45.

[0004] In the document-retrieval equipment of the above-mentioned composition, after an input keyword and significance are acquired, process until two or more keywords used for thesaurus reference are specified is shown in drawing 5. Signs 42, 47, and 49 express applicable functional block of the keyword determination equipment 41 in drawing 4 among drawing 5.

[0005] Reference of drawing 5 **** "kwd8", "kwd6", and "kwd7" from the thesaurus dictionary 42, respectively as an additional keyword relevant to an input keyword "kwd1", "kwd2", and "kwd3." Moreover, "2", "10", and "6" are set up, respectively as a significance of an input keyword "kwd1", "kwd2", and "kwd3." About the significance of an additional keyword "kwd8", "kwd6", and "kwd7", "one half of the values of the value of the significance with which these were set as an input keyword "kwd1", "kwd2", and "kwd3" in consideration of being led secondarily from the input keyword, "1", "5", and "3" are set up. [i.e.,] DB management equipment 44 performs thesaurus reference (document retrieval) using the keyword and significance of the above-mentioned plurality, **** the corresponding electronic filing document from a document DB43, and sends this to a display controller 45. A display controller 45 carries out a screen display of this electronic filing document to display 46.

[0006] Thus, with conventional document-retrieval equipment, by performing thesaurus reference by eight keywords which added not only "kwd4" but four input keyword "kwd1" additional keywords "kwd6" - "kwd9", while enabling wide range reference, acquisition of the electronic filing document which agreed for the purpose of reference by taking the significance of each keyword into consideration at the time of reference is enabled.

[0007]

[Problem(s) to be Solved by the Invention] As mentioned above, with conventional document-retrieval

equipment, all additional keywords are used as a keyword for thesaurus reference. However, into an additional keyword, it is not necessarily important for an operator, or a completely unnecessary thing may be contained and the thesaurus reference based on such an unnecessary additional keyword serves as useless processing. Therefore, with conventional document-retrieval equipment, an electronic filing document which it not only cannot attain shortening of retrieval time, but is completely different from the electronic filing document which an operator desires truly was ****(ed) in many cases. For example, when "Mac" who is the abbreviation of the "Macintosh" which is the product name of a personal computer is made into an input keyword, "Macdonald" who is the chain store of a fast food may be extracted as an additional keyword. In this case, the electronic filing document which described the food industry-related information which is completely unrelated to a computer will be searched.

[0008] Then, in the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible, in case the technical problem of this invention adds other related keywords to an input keyword, it is for the related keyword to judge whether it is a keyword very required for reference, and avoid useless reference processing for an operator. Other technical problems of this invention are to offer the record medium for a computer realizing the above-mentioned document-retrieval processing.

[0009]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, this invention offers the improved keyword determination method which is applied in the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible. When this method makes the electronic description information of a keyword candidate and the keyword candidate concerned link to document-retrieval equipment beforehand, and is held and arbitrary input keywords are received, on document-retrieval equipment The processing which acquires the electronic description information linked to other keyword candidate who should add to the input keyword concerned, and its keyword candidate, The processing which presents the keyword candidate who acquired with electronic description information with the gestalt which an operator can recognize, It is characterized by performing at least processing which determines 1 or two or more keywords which are used for thesaurus reference based on the additional propriety information of the keyword candidate concerned that it was inputted by the operator according to this presentation in this order. Electronic description information is the description document which recorded the circumference knowledge which contains the keyword candidate, for example, and when it is shown, it is the thing of the content as which an operator can grasp a keyword candidate's semantic content.

[0010] Thus, as facilities are given to judgment whether a keyword candidate should be added to an operator by showing an operator a keyword candidate and electronic description information collectively from a document-retrieval equipment side and the additional propriety of a keyword is determined according to the judgment result by the operator, the reference processing based on an unnecessary keyword is avoided.

[0011] this invention also offers the document-retrieval equipment suitable for operation of the above-mentioned method again. The document-retrieval equipment of this invention is equipment in which the thesaurus reference by two or more keywords related mutually is possible, and is constituted including the keyword determination equipment which determines 1 or two or more keywords which uses for thesaurus reference based on the recording information of the thesaurus dictionary which recorded the keyword candidate who should add to each keyword, the information file which recorded the keyword candidate's electronic description information for every keyword candidate, and the aforementioned thesaurus dictionary and an information file.

[0012] Keyword determination equipment consists of all or a part of combination of the following components, and it determines the keyword for performing rational reference based on the additional propriety information from an operator while making the presentation gestalt which urges the input of the aforementioned keyword candidate's additional propriety information to document-retrieval equipment to an operator form.

(1) A keyword information acquisition means to acquire the electronic description information of the keyword candidate who should add to an input keyword, and its keyword candidate.

(2) A score grant means to give the score which expresses a related degree with the aforementioned input keyword to the keyword candidate or electronic description information acquired with the keyword information acquisition means.

(3) Presentation control means to which the keyword candidate who acquired with the aforementioned keyword information acquisition means is made to show with the electronic description information with the gestalt which an operator can recognize. In addition, when a score is given to electronic description information as mentioned above, these presentation control means are constituted so that a keyword

candidate and electronic description information may be made to show with a different gestalt according to a score.

(4) A means to determine whether to decode a keyword candidate's additional propriety information that it was inputted by the operator according to presentation by the aforementioned presentation control means, and add the keyword candidate concerned as a keyword for thesaurus reference according to a decode result.

(5) A means to add the weight information according to the reference purpose to each of the determined keyword for thesaurus reference.

[0013] The record medium of this invention which solves a technical problem besides the above It is the record medium with which the program for realizing on a computer the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible was recorded with the gestalt in which computer reading is possible. When the processing and the arbitrary input keywords which the aforementioned program makes link the electronic description information of a keyword candidate and the keyword candidate concerned beforehand, and holds at least are received The processing which acquires the electronic description information linked to other keyword candidate who should add to the input keyword concerned, and its keyword candidate, The processing which presents the keyword candidate who acquired with electronic description information with the gestalt which an operator can recognize, The aforementioned computer is made to perform processing which determines 1 or two or more keywords which are used for thesaurus reference based on the additional propriety information of the keyword candidate concerned that it was inputted by the operator according to this presentation.

[0014] Other record media of this invention are record media with which the program for realizing on a computer the document-retrieval equipment in which the thesaurus reference by two or more keywords related mutually is possible was recorded with the gestalt in which computer reading is possible. The information file which recorded the keyword candidate's electronic description information for every thesaurus dictionary in which the aforementioned program recorded at least the keyword candidate who should add to each keyword, and keyword candidate is made to form in the aforementioned computer. further While acquiring the keyword candidate who should add to an input keyword from the aforementioned thesaurus dictionary The processing which acquires the electronic description information corresponding to this keyword candidate from the aforementioned information file, The processing which presents the keyword candidate and electronic description information which were acquired with the gestalt which an operator can recognize, A keyword candidate's additional propriety information that it was inputted by the operator according to this presentation is decoded, and the aforementioned computer is made to perform processing which determines whether add the keyword candidate concerned as a keyword for thesaurus reference according to a decode result.

[0015]

[Embodiments of the Invention] Hereafter, with reference to a drawing, the form of operation of this invention is explained in detail. Drawing 1 is the functional block diagram of the document-retrieval equipment concerning 1 operation form of this invention. This document-retrieval equipment is realized by computers, such as a personal computer and a workstation. Namely, it installs in the computer memory which does not illustrate the program which it was inputted through the communication network or was read from portability media, such as CD-ROM (compact disk type ROM) and FD (flexible disk). The keyword determination equipment 1 formed when this computer is beginning to read the above-mentioned program at any time and executes it, the thesaurus dictionary 2, the description document file 3 which is an information file, a document DB4, DB management equipment 5, and a display controller 6, It is constituted including the display 7 which consists of CRT or a liquid crystal panel.

[0016] The thesaurus dictionary 3, a document DB4, DB management equipment 5, and display 7 are the things of the same function as the thesaurus dictionary 42 of equipment, a document DB43, DB management equipment 44, and display 46 conventionally which was shown in drawing 4, respectively. The description document (electronic description information) which described the keyword candidate's semantic content, circumference knowledge, other related information, etc. for every keyword candidate is recorded on the description document file 3. In this case, the description document prepares two or more things about one keyword candidate, and enables it to choose a specific thing according to each related degree.

[0017] Keyword determination equipment 1 is constituted in the function of the data input section 11, the retrieval-by-keyword section 12, the file management section 13, the score grant section 14, and the keyword specification section 15, in order to determine the rational keyword used for thesaurus reference. The data input section 11 receives the entry of data of an input keyword, its significance, a keyword

candidate's additional propriety information, and others, and holds the retrieval-by-keyword section 12 temporarily as a keyword candidate who **** other keywords relevant to an input keyword from the thesaurus dictionary 2, and should add this.

[0018] The file management section 13 makes the record section of the description document of a keyword candidate and its keyword candidate link, and is managed. It has the table which specifically means whether a certain keyword candidate's description document exists in the record section (code number) of description document file 3 throat, and when a keyword candidate is specified, the description document about this keyword candidate can be acquired from the description document file 3 through a code number.

[0019] The score grant section 14 gives the score (evaluation value) which expresses a related degree with an input keyword or a keyword candidate to each description document. This score is used at the time of specification of selection sequence in case there are two or more description documents about the same keyword candidate, and formation of the display form for making easy an additional propriety judgment of the keyword candidate by the operator in a display controller 6. About grant of a score, various technique is employable. For example, the well-known TF/IDF method ("Introduction to Modern Information Retrieval", work besides Gerard Salton, McGraw-Hill Publishing Company **), the score grant technique currently indicated by the Japanese Patent Application No. 110870 [eight to] specification "similar document-retrieval equipment" previously proposed by these people can be used.

[0020] When the additional propriety information of the keyword candidate concerned that it was inputted by the operator is decoded and it judges with additional needlessness (an operator does not employ the keyword candidate) as a result of decode, the keyword specification section 15 cancels the keyword candidate (deletion), and makes only an input keyword the keyword for reference. On the other hand, when an addition judges with the need (an operator employs the keyword candidate) as a result of decode, the keyword candidate is added to an input keyword. Moreover, the significance of an input keyword and the added keyword is also determined, these are set up for every keyword, and it outputs to DB management equipment 5. What was inputted with the keyword is used for the significance of an input keyword as it is. About the significance of the added keyword, the value of the significance of an input keyword may be used as it is, and the value which carried out the multiplication of the coefficient (for example, the 0.5 [conventionally same] as the case of equipment) to the value of the significance of an input keyword may be used. The latter value is used in this example.

[0021] A display controller 6 receives the description document and keyword candidate to whom the score was given from keyword determination equipment 1, and performs control for displaying a keyword candidate and its description document on display 7 simultaneously. The form to display enables it to choose arbitrarily various forms to which an operator can call attention, such as a form on which some of for example, keyword candidates and description documents (for example, only a title chisel, a title and an author, or a summary) are displayed, a form on which the whole sentence of a keyword candidate and a description document is displayed, and a form to which highlighting only of the keyword candidate is carried out using a different color from blink or other portions. Moreover, in order that an operator may enable it to judge a keyword candidate's additional propriety easily, the above-mentioned score given to the description document is discriminated, and you may locate the higher description document of a score in the screen upper part, or may be made to carry out highlighting in a big character, the strong color of impact, etc. Moreover, a description document is displayed in stack form and you may make it display the higher description document of a score on the high order of a stack.

[0022] Next, with reference to drawing 2 and drawing 3, operation of the document-retrieval equipment by this operation form is explained concretely. Drawing 2 is the input keyword and keyword candidate who use in the document-retrieval equipment of this operation form, a description document, a score, and explanatory drawing showing the example of significance, and drawing 3 is a flow chart which mainly shows the procedure of the keyword specification processing in keyword determination equipment 1. In drawing 2, signs 2, 3, 7, 11, 13, and 15 express applicable functional block of the keyword determination equipment 1 in drawing 1, respectively.

[0023] Suppose that three input keywords, "AIDS", "phytotoxicity", and "infection", were inputted into the data input section 11 as now shown in drawing 2 (Step S101). The retrieval-by-keyword section 12 searches whether the keyword candidate who should add to the thesaurus dictionary 2 about "AIDS" is recorded first (Step S102). In this example, since the "blood product" is recorded on the thesaurus dictionary 2 as a keyword candidate to "AIDS", the retrieval-by-keyword section 12 **** this "blood product", and sends it to the file management section 13.

[0024] The file management section 13 specifies the code number "4000" of the field where it is recorded from the "blood product" which is a keyword candidate. And description document "1" - "3" is ****(ed) from the record section of the description document file 3 corresponding to this code number "4000", and

this is sent to the score grant section 14 (Step S103).

[0025] The score grant section 14 gives a score to description document "1" - "3" based on an input keyword, its significance, and a keyword candidate, respectively (Step S104). Since the "AIDS" of significance "2", the "phytotoxicity" of significance "10", and the "infection" of significance "6" are included in the description document "1", the score is higher than other description documents "2" and "3." Description document "1" - "3" to which the score was given is sent to a display controller 6 with the "blood product" which is a keyword candidate. Both the display controllers 6 display "blood product" and each description document "1" - "3" on display 7 (Step S105).

[0026] If a keyword candidate's additional propriety information is inputted from an operator based on the display of a "blood product" etc., the keyword specification section 15 will decode this additional propriety information. And when addition of a "blood product" is not chosen as a result of decode (Step S106: No), the keyword specification section 15 deletes the "blood product" currently held at self or the retrieval-by-keyword section 12. On the other hand, when addition of a "blood product" is chosen (Step S106: Yes), the significance of a "blood product" is set up (Step S108). With this operation form, significance 0.5 times the value of the input keyword of a basis is made into a keyword candidate's significance, and since the significance of an input keyword "AIDS" is "2", the significance of a "blood product" is set as "1."

[0027] When there is the next keyword candidate, it returns to processing of Step S102, and when there is no next keyword candidate, the keyword for reference is determined and it outputs to DB management equipment 5 (Step S109). That is, when an operator does not choose addition of a "blood product", only three input keywords, "AIDS", "phytotoxicity", and "infection", are determined as a keyword for reference, when addition of a "blood product" is chosen on the other hand, this is added to the keyword for reference, and it outputs to DB management equipment 5. DB management equipment 5 searches an electronic filing document, after also taking such significance into consideration for the determined keyword for reference.

[0028] Thus, with the document-retrieval equipment of this operation form, since the keyword for reference is determined after making an operator judge the additional propriety of the keyword candidate by whom keyword addition equipment 1 was ****(ed) based on the input keyword, the reference processing by the unnecessary keyword is avoidable. Moreover, since it was made to also display the description document with a keyword candidate, it can judge now whether even if an operator does not investigate separately a keyword candidate's semantic content or its related knowledge, the keyword candidate is exactly [easily and] required for reference, and shortening of the time which reference takes can also be attained. Furthermore, since thesaurus reference can be performed after also taking into consideration the significance given to each keyword, reference precision can also be raised.

[0029] In addition, although this operation form explained the example in the case of processing every one keyword candidate, you may be made to perform the above-mentioned processing simultaneously about two or more keyword candidates. Moreover, with this operation form, although the number of keyword candidates to one input keyword was explained as one, the number of keyword candidates to an input keyword can be determined arbitrarily.

[0030] Furthermore, although the example of the screen display using display 6 as a form of presentation of the keyword candidate to an operator or electronic description information was explained, it replaces with a screen display or you may make it show in an audio form with a screen display in this operation form using a voice synthesizer, the existing text reader, etc. When outputting a keyword candidate etc. with voice, an above-mentioned score can be used as a parameter for changing the strength and the tone of volume.

[0031]

[Effect of the Invention] since an operator is shown a keyword candidate with the electronic description information according to this invention so that clearly from the above explanation, an operator is easy in judgment whether an addition of the keyword candidate is required -- and it can carry out now correctly

[0032] Moreover, only the keyword candidate who judged that it was required is added by the operator, since an excessive keyword candidate will not be used for reference, useless reference processing is avoided and the much more high-speed and exact document retrieval of him becomes possible from the former.

[0033] Furthermore, since electronic description information is given and shown the score according to the related degree with an input keyword, additional propriety judgment of the keyword candidate by the operator becomes still more easily and exact.

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The functional block diagram of the document-retrieval equipment concerning 1 operation gestalt of this invention.

[Drawing 2] Explanatory drawing showing the example of the input keyword in the document-retrieval equipment of this operation gestalt, a keyword candidate, a description document, a score, and significance, and process until the keyword for reference is specified based on these.

[Drawing 3] The flow chart which shows the procedure of the keyword specification processing by the document-retrieval equipment of this operation gestalt.

[Drawing 4] The functional block diagram of conventional document-retrieval equipment.

[Drawing 5] Explanatory drawing showing the example of the input keyword in conventional document-retrieval equipment, an additional keyword, and significance, and process until the keyword for reference is specified.

[Description of Notations]

1 Keyword Determination Equipment

2 Thesaurus Dictionary

3 Description Document File

4 Document DB (Database)

5 DB Management Equipment

6 Display Controller

7 Display

11 Data Input Section

12 Retrieval-by-Keyword Section

13 File Management Section

14 Score Grant Section

15 Keyword Specification Section

[Translation done.]